Applied Mathematics Letters 23 (2010) 605–608

Contents lists available at ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

Robust least squares solution of linear inequalities

Maziar Salahi

Department of Mathematics, Faculty of Sciences, University of Guilan, P.O. Box 1914, Rasht, Iran

ARTICLE INFO

Article history:
Received 3 April 2009
Received in revised form 8 January 2010
Accepted 27 January 2010

Keywords:
Linear inequalities
Least squares solution
Second order cone
Interior point methods

ABSTRACT

Least squares solution of linear inequalities appears in many disciplines such as linear separability problems and inconsistency correction. In this paper we consider this problem with uncertainty in its data. Then we prove that its robust counterpart is equivalent to a second order conic linear optimization problem, which can be efficiently solved using interior point methods.

1. Introduction

Suppose a set of observations places upper or lower bounds on linear combinations of some variables, we want to find x^* the solution of

$$\min_{x\in R^n} \|(Ax-b)_+\|^2 \tag{1}$$

where $a_+ = \max(a, 0)$. This is known as the least squares solution of linear inequalities and is applicable in areas like linear separability problems, inconsistency correction and many others [1–4]. For example, in linear separability problems, the goal is to find a hyperplane that best separates two point sets; when the two sets are not linearly separable, a hyperplane that correctly separates the largest number of points is desired. In the Euclidean norm, finding such a hyperplane requires solving problem (1).

As is known [1] the objective function of (1) is a once differentiable convex function and the necessary and sufficient condition for x^* to be an optimal solution is $A^T(Ax^*-b)_+ = 0$. Using this condition, an efficient algorithm has been designed in [1]. Although the objective function in (1) is a once differentiable convex function, using the notion of generalized Hessian an efficient second order algorithm is also introduced in [4] to solve (1). However, the uncertainty of the problem data is not taken into account in the existing algorithms, while b or A, or both of them might be uncertain. In such a case, problem (1) with extra uncertainty constraints cannot be handled by the classical algorithms [1,4]. In this paper, we show that the robust counterpart of the problem, when both A and b are contaminated by noise, is equivalent to a second order conic linear program, which can be efficiently solved using interior point methods [5,6]. Furthermore, the structured robust counter part of a model which arises in linear separability problems is given as a second order conic linear program. Throughout this paper we use Euclidean norms unless clearly stated.

2. Robust optimization approach

In this section first let us consider the case where the right hand side vector in (1) is uncertain. Further suppose we allow the amount of uncertainty in this vector to be bounded by for example $\rho > 0$. Then for a given $x \in R^n$, the worst case

E-mail addresses: salahim@guilan.ac.ir, salahi.maziar@gmail.com.

0893-9659/$ – see front matter
doi:10.1016/j.aml.2010.01.020

infeasibility is

$$\max_{\|r\|\leq\rho} \|(Ax-(b+r))_+\|. \tag{2}$$

Now a robust solution is the one which minimizes this worst case infeasibility i.e.,

$$f(A,b,\rho) := \min_{x\in R^n}\max_{\|r\|\leq\rho} \|(Ax-(b+r))_+\|. \tag{3}$$

Obviously $f(A,b,\rho)=\rho f\left(\frac{A}{\rho},\frac{b}{\rho},1\right)$. Thus for simplicity in the rest of the paper we consider $\rho=1$.

Theorem 2.1. *Problem (3) with $\rho=1$ is equivalent to*

$$\min_{x\in R^n} \|(Ax-b)_+\|+1. \tag{4}$$

Proof. For a fixed $x\in R^n$ we have

$$\max_{\|r\|\leq 1} \|(Ax-(b+r))_+\| \leq \|(Ax-b)_+\|+1.$$

Now if we choose $r=-\frac{(Ax-b)_+}{\|(Ax-b)_+\|}$ then the equality holds. Therefore, (3) is equivalent to

$$\min_{x\in R^n} \|(Ax-b)_+\|+1. \quad \square$$

Remark 2.2. As we see, when the vector b is uncertain, the optimal solution is the same as (1). It should be noted here that if A is ill-conditioned, then special care should be taken in solving (4) as it might give solutions with very large norm and meaningless from practical point of view.

Now let us consider the case where both A and b are uncertain and the amount of uncertainty in the Frobenius norm is bounded by one. Then minimizing the worst case infeasibility is

$$\min_{x\in R^n}\max_{\|[E\ r]\|_F\leq 1} \|((A+E)x-(b+r))_+\|. \tag{5}$$

To see an equivalent reformulation of (5) we need to introduce the second order cone [7,8].

Definition 2.3. A second order (Lorentz) cone in R^n is denoted by Q_n and is defined as

$$Q_n=\left\{(x_1,x_2,\ldots,x_n)\in R^n \middle| \sqrt{x_2^2+\cdots+x_n^2}\leq x_1\right\}.$$

Analogous to the nonnegative orthant as a cone, Q_n has the following fundamental properties that are crucial in developing interior point algorithms for solving second order conic optimization problems:

- Q_n is a closed and convex cone.
- Q_n is self dual.
- Q_n is pointed and has nonempty interior.

Let us denote by K the product of nonnegative orthant and second order cones. A primal standard form for a linear conic optimization problem is given by

$$\begin{array}{ll} \min & c^T x \\ & Ax=b \\ & x\in K \end{array} \tag{6}$$

and its dual is

$$\begin{array}{ll} \max & b^T y \\ & c-A^T y\in K. \end{array} \tag{7}$$

As we see, if we simply let $K=R^n_+$, then we have a linear programming with its dual. However, in linear conic optimization framework, K is a product of several linear and second order cones with different sizes.

Theorem 2.4. *Problem (5) is equivalent to the following conic-linear minimization problem:*

$$\begin{array}{ll} \min & t+s \\ & (Ax-b)\leq y \\ & \begin{pmatrix} t \\ y \end{pmatrix}\in Q_{m+1}, \qquad \begin{pmatrix} s \\ 1 \\ x \end{pmatrix}\in Q_{n+2} \\ & y\geq 0. \end{array} \tag{8}$$

Proof. For a fixed $x \in R^n$ we have

$$\max_{\|[E\ r]\|_F \le 1} \|((A+E)x-(b+r))_+\| \le \|(Ax-b)_+\| + \sqrt{1+\|x\|^2}.$$

Now if we choose $E = \frac{(Ax-b)_+ x^T}{\sqrt{1+\|x\|^2}\,\|(Ax-b)_+\|}$, $r = -\frac{(Ax-b)_+}{\sqrt{1+\|x\|^2}\,\|(Ax-b)_+\|}$, then $\|[E\ r]\|_F = 1$ and

$$\|((A+E)x-(b+r))_+\| = \|(Ax-b)_+\| + \sqrt{1+\|x\|^2}.$$

Thus (5) is equivalent to

$$\min_{x\in R^n} \|(Ax-b)_+\| + \sqrt{1+\|x\|^2}.$$

This further can be presented in the linear conic form as

$$\begin{aligned} \min\quad & t+s \\ & \|(Ax-b)_+\| \le t \\ & \sqrt{1+\|x\|^2} \le s \end{aligned}$$

or

$$\begin{aligned} \min\quad & t+s \\ & (Ax-b) \le y \\ & \|y\| \le t \\ & \sqrt{1+\|x\|^2} \le s \\ & y \ge 0 \end{aligned}$$

or

$$\begin{aligned} \min\quad & t+s \\ & (Ax-b) \le y \\ & \begin{pmatrix} t \\ y \end{pmatrix} \in Q_{m+1}, \qquad \begin{pmatrix} s \\ 1 \\ x \end{pmatrix} \in Q_{n+2} \\ & y \ge 0. \quad \square \end{aligned}$$

Remark 2.5. Obviously (8) is in the dual form (7) and one can use existing efficient software like Mosek and SeDuMi [5,6] to solve it.

In the sequel we give a second order conic linear programming model for the robust linear separability problem. Suppose $\mathcal{A}$ and $\mathcal{B}$ are two sets in R^n with m and k points, respectively. Now we want to find a hyperplane $w^T x = \gamma$ that is optimal in the sense of having fewest points incorrectly classified as belonging to $\mathcal{A}$ or $\mathcal{B}$. This can be approximated as the least squares solution of

$$\begin{aligned} & Aw - \gamma e_m \le -e_m \\ & -Bw + \gamma e_k \le -e_k \end{aligned} \tag{9}$$

given by

$$\min \|G\hat{w} - g\| \tag{10}$$

where

$$G = \begin{bmatrix} A & -e_m \\ -B & e_k \end{bmatrix}, \qquad g = \begin{bmatrix} -e_m \\ -e_k \end{bmatrix}, \qquad \hat{w} = \begin{bmatrix} w \\ \gamma \end{bmatrix}.$$

However, there might be uncertainty in our data sets $\mathcal{A}$ and $\mathcal{B}$. As we see, in such a case, the whole matrix G is not uncertain, in other words there are structured uncertainty in it. In the next theorem we give a conic reformulation of robust counterpart of (10) with the one as the level of the uncertainty, namely

$$\min_{\hat{w}} \max_{\|E\|_F \le 1} \|((G+E)\hat{w} - g)_+\|, \tag{11}$$

where

$$E = \begin{bmatrix} E_1 & 0 \\ E_2 & 0 \end{bmatrix}.$$

Theorem 2.6. *Problem (11) is equivalent to*

$$\begin{aligned} \min \quad & t+s \\ & G\hat{w}-g \leq y \\ & \|w\| \leq s \\ & \|y\| \leq t \\ & y \geq 0. \end{aligned} \tag{12}$$

Proof. We have

$$\|((G+E)\hat{w}-g)_+\| \leq \|(G\hat{w}-g)_+\| + \|(E\hat{w})_+\| \leq \|(G\hat{w}-g)_+\| + \|w\|.$$

Now if we choose

$$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \frac{(G\hat{w}-g)_+ w^T}{\|(G\hat{w}-g)_+\| \, \|w\|},$$

then in the previous inequality equality holds. Thus (11) is equivalent to

$$\min_{\hat{w}} \|(G\hat{w}-g)_+\| + \|w\|.$$

This further can be written as the second order conic linear program (12). □

3. Conclusions

In this paper we have proved that the robust counterpart of the least squares solution of linear inequalities is equivalent to a second order conic linear program which is efficiently solvable by interior point method based software packages.

References

[1] R. Bramley, B. Winnicka, Solving linear inequalities in a least squares sense, SIAM Journal on Scientific Computing 17 (1) (1996) 275–286.
[2] Y. Censor, A. Ben-Israel, Y. Xiao, J.M. Galvin, On linear infeasibility arsing in intensity-modulated radiation teraphy inverse planning, Linear Algebra and its Applications 428 (2008) 406–1420.
[3] A.I. Golikov, Yu.G. Evtushenko, S. Ketabchi, On families of hyperplanes that separate polyhedra, Computational Mathematics and Mathematical Physics 45 (2) (2005) 227–242.
[4] M. Salahi, S. Ketabchi, Correcting an inconsistent set of linear inequalities by generalized Newton method, Optimization Methods and Software (2010) (in press).
[5] E.D. Andersen, K.D. Andersen, The MOSEK interior point optimizer for linear programming: An implementation of the homogeneous algorithm, in: H. Frenk, K. Roos, T. Terlaky, S. Zhang (Eds.), High Performance Optimization, Kluwer Academic Publishers, 2000, pp. 197–232.
[6] J.F. Sturm, Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optimization Methods and Software 11–12 (1999) 625–653.
[7] F. Alizadeh, D. Goldfarb, Second order cone programming, Mathematical Programming, Series B 95 (2003) 3–51.
[8] A. Ben-Tal, A. Nemirovski, Lectures on Modern Convex Optimization, in: AMS–SIAM Series on Optimization, SIAM, Philadelphia, 2002.